Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 5655–5659, June 1992
Genetics

The *patA* gene product, which contains a region similar to CheY of *Escherichia coli*, controls heterocyst pattern formation in the cyanobacterium *Anabaena* 7120

(differentiation/signal transduction/regulatory circuit)

JIHONG LIANG, LORI SCAPPINO, AND ROBERT HASELKORN*

Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637

Contributed by Robert Haselkorn, March 18, 1992

ABSTRACT **In *Anabaena* 7120, heterocysts (cells specialized for nitrogen fixation) develop at the ends of filaments and at intervals within each filament. We have isolated a mutant *Anabaena* strain that develops heterocysts mostly at the ends of filaments. This mutant, PAT-1, grows poorly under nitrogen-fixing conditions. The wild-type gene that complements the mutation in PAT-1, called *patA*, was cloned and sequenced. The predicted PatA protein contains 379 amino acids distributed among three "domains" based on predictions of hydropathy and flexibility. The carboxyl-terminal domain is very similar to that of CheY and other response regulators in two-component regulatory systems in eubacteria. The *patA* mutation suppresses the multiheterocyst phenotype produced by extra copies of the wild-type *hetR* gene described previously, suggesting that PatA and HetR are components of the same environment-sensing regulatory circuit in *Anabaena*.**

Anabaena 7120 is a member of the class of filamentous cyanobacteria that are capable of both photosynthesis and nitrogen fixation under aerobic conditions. In the presence of a source of combined nitrogen, *Anabaena* 7120 grows in continuous chains of vegetative cells. When the filaments are deprived of fixed nitrogen, some cells differentiate into thick-walled cells called heterocysts that become capable of nitrogen fixation (1). The heterocyst provides a strictly anaerobic environment that is required by the nitrogen-fixation machinery (2).

In the laboratory, *Anabaena* 7120 differentiates heterocysts both terminally and internally at intervals of ≈10 cells, although the intervals are slightly longer in older cultures. Division of vegetative cells widens the space between heterocysts, but the spacing pattern is maintained by the formation of additional heterocysts midway between two existing heterocysts (3, 4). The placement of heterocysts can be altered by physical breakage of the filaments, by modulation of the light intensity, or by treatment with chemicals such as 7-azatryptophan (5) or rifampicin (4). Apparently, many, if not all, of the cells in a filament have the potential to develop into heterocysts. The mechanism that allows cells only at certain positions in the filament to differentiate into additional heterocysts must involve intercellular communication. It has been suggested that the establishment and maintenance of the heterocyst spacing pattern depend on (*i*) the interaction between diffusible substances that originate in the heterocyst and move along the filament and (*ii*) other factors, presumably proteins, that both interact with the diffusible substances and control, directly or indirectly, gene expression (6–8).

Mutations affecting the heterocyst pattern have been identified in several *Anabaena* strains (9, 10). In one case, the mutant produces no heterocysts under inducing conditions (10). The wild-type gene complementing this mutation, called *hetR*, was cloned and sequenced (11). When an extra copy of the wild-type *hetR* gene carried on a plasmid was introduced into wild-type cells, the heterocyst frequency was increased due to the induction of heterocysts in strings of two, three, or more cells (11). The expression of *hetR* occurs only in the cells that are going to differentiate into heterocysts, as seen by chemiluminescence in cells containing a *lux* gene fused to the upstream region of the *hetR* gene (12).

Here, we describe another mutant of *Anabaena* 7120, called PAT-1, that differentiates heterocysts mostly at the ends of filaments. The sequence of the gene *patA* that complements the mutation in PAT-1 shows it to be related to *cheY*, a component of an environment-sensing regulatory circuit in *Escherichia coli* (13, 14). The *patA* mutation suppresses the multiheterocyst phenotype of the extracopy *hetR* strain.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions. *E. coli* strains MC1061 and DH5α were used as the hosts for plasmid preparation and for conjugal transfer of DNA to *Anabaena* (15). For selective growth, Luria broth was supplemented with antibiotics at the following concentrations: ampicillin at 100 μg/ml, kanamycin at 50 μg/ml, chloramphenicol at 10 μg/ml, and spectinomycin at 20 μg/ml. *Anabaena* 7120 and its mutants were grown in BG-11 medium (16). For plate culture, the BG-11 liquid medium was supplemented with 10 mM Hepes buffer, pH 8.0/1 mM NaS_2O_3/1.5% agar (BBL purified; Becton Dickinson). For selective growth, antibiotics were added to final concentrations of 30 μg/ml (neomycin), 10 μg/ml (chloramphenicol), and 5 μg/ml (erythromycin). A combination of spectinomycin and streptomycin was used at a final concentration of 2 μg/ml each.

Mutagenesis and Isolation of Mutants. Tn*5* mutagenesis was done as described (17). *E. coli* strain MC1061 was successively transformed with plasmids pRK24, pRL528, and pBR322::Tn*5*. Equal volumes of *E. coli* and *Anabaena* cell suspensions were mixed and spread on a nitrocellulose membrane placed on BG-11-agar plates, grown for 3 days, at which time the membranes were transferred to other BG-11 plates containing neomycin (30 μg/ml). When well-separated green colonies appeared (≈2 weeks), the membranes were transferred to BG-11 N^- plates. Colonies that turned yellow around the edges (Fix^- phenotype) were picked and maintained on BG-11 plates.

Complementation of Mutant PAT-1. Complementation of mutant PAT-1 with the cosmid bank containing fragments of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: ORF, open reading frame.

*To whom reprint requests should be addressed at: Department of Molecular Genetics and Cell Biology, University of Chicago, 920 East 58th Street, Chicago, IL 60637.

wild-type DNA was conducted essentially as described (10). A similar procedure was used for subsequent confirmatory complementation. For this procedure, the *E. coli* donor carried plasmids pRK24, pRL528, and the complementing plasmid. After 3-day incubation on a BG11-agar plate, the membrane was transferred to a BG11-N^- plate containing neomycin at 30 μg/ml. Successful complementation gives rise to a green cell lawn, whereas the mutant culture exhibits a yellow-green color.

DNA Subcloning. The DNA of a cosmid that complemented mutant PAT-1 was partially digested with *Sau3*AI and fractionated by agarose gel electrophoresis. Fragments of 1–2, 2–4, and 4–14 kilobases (kb) were isolated with a GeneClean kit (Bio 101, La Jolla, CA) and separately ligated into the *Bam*HI site of pCCB110 (a derivative of the shuttle vector pRL25c with the polylinker of pIC20R inserted into the *Eco*RI site; C. Bauer, personal communication). The plasmid banks were subsequently conjugated into *Anabaena* mutant PAT-1 for the complementation test.

DNA Sequence Analysis. Plasmids pJHLb-5 and pJHLb-10 complement mutant PAT-1 and contain the same 2.7-kb DNA fragment inserted in the opposite orientation in the vector pCCB110. Nested deletions were constructed by using an Erase-a-Base kit from Promega. Complete sequences of both strands were determined by using the sequencing kit from United States Biochemical and are available from GenBank under accession number M87501. Sequences were analyzed with the University of Wisconsin Genetic Computer Group's software package and programs written by William Buikema (University of Chicago) and Conrad Halling (University of Chicago) for the Apple Macintosh.

Insertional Mutation of the *patA* Gene in *Anabaena*. The 2.7-kb fragment of *Anabaena* DNA in pJHLb-5 was subcloned into *Eco*RI/*Pst* I-digested pUC18 to create pJHL105. The Spc^r/Sm^r Ω cassette from pDW9 (18) was placed into the *Hin*dIII site in *patA* in pJHL105 to produce pJHL106. The entire 2.7-kb fragment with the inserted Ω cartridge was then isolated from pJHL106 by *Pst* I and *Sst* I digestion and ligated into pRL271 (19), which carries Cm^r and Em^r genes and the *sacB* gene of *Bacillus subtilis*, treated with the same enzymes. The resultant plasmid, pJHL107, was mobilized into *Anabaena* 7120 by conjugation as described above. Exconjugants were selected on BG-11 plates containing spectinomycin and streptomycin. Double recombinants that contain only the interrupted copy of the *patA* gene were obtained by screening 100 Spc^r/Sm^r colonies on BG-11 plates containing Cm (10 μg/ml) and Em (5 μg/ml). Southern hybridization to genomic DNA digests of both single and double recombinants confirmed these constructions.

PCR Amplification of the Mutant Copy of the *patA* Gene. A 1.8-kb fragment containing the entire *patA* gene from the mutant strain PAT-1 was amplified by using the techniques recommended by the supplier of the GeneAmp kit (Perkin–Elmer/Cetus). In making primers, five nucleotides were added to the 5′ end of each primer to create a *Bam*HI or *Bgl* II restriction site for the convenience of subsequent cloning. The 1.8-kb PCR fragment was cut into five short sections with *Sau3*AI and cloned into the *Bam*HI site of pUC18. Sequences of these fragments were determined as described above.

Microscopy. Morphological examination and photomicroscopy of *Anabaena* cells were conducted with a Zeiss Axioskop microscope equipped with differential-interference-contrast (Nomarski) and phase-contrast optics.

RESULTS

Isolation of Mutants Altered in Heterocyst Spacing Pattern. We tried to introduce the transposon Tn*5* carried on pBR322 into wild-type *Anabaena* 7120 cells by using the conjugation system developed by Elhai, Wolk, and coworkers (20, 21). Exconjugants that were resistant to neomycin (30 μg/ml) were screened for their ability to grow on N^- plates. Colonies that gradually turned yellow around the edges, the expected Fix^- phenotype, were picked and maintained on N^+ plates. These Fix^- mutants were examined under the microscope. Two mutants, among 89 examined, exhibited altered heterocyst-spacing patterns. One mutant forms many more heterocysts than the wild type, whereas the other mutant forms fewer heterocysts. The latter is the subject of this report; studies of the other mutant will be reported elsewhere.

Although the mutant, called PAT-1, was originally picked from neomycin-resistant colonies, it became neomycin-sensitive subsequently. This neomycin sensitivity could be caused by a mutation in the gene coding for the antibiotic resistance or could be from the loss of Tn*5* from the chromosome. To test these possibilities, we used an internal fragment of the *npt* gene of Tn*5* to probe *Hin*dIII- and/or *Eco*RI-digested chromosomal DNA of both wild-type and the mutant PAT-1. No signal was detected by autoradiography (data not shown). The lack of a Tn*5* insertion in the chromosome of the mutant strain was further confirmed by Southern hybridization of the *patA* gene to *Hin*dIII-digested chromosomal DNA (data not shown), which showed no difference in the size of the *patA*-containing bands in either strain. Thus, the mutation in PAT-1 either is spontaneous or is due to the insertion and subsequent excision of Tn*5*. DNA sequencing (see below) suggests the mutation to be the former.

Strain PAT-1 grows as well as the wild type in medium containing either NH_4Cl or $NaNO_3$. Under these conditions, the two strains are morphologically identical. In the absence of fixed nitrogen, mutant PAT-1 grows poorly, presumably due to the low frequency of heterocysts (Fig. 1), although the specific activity of nitrogenase, as measured by acetylene-reduction assay, is comparable to that of the wild type (data not shown). Upon induction, heterocysts initially differentiate at the ends of filaments. The formation of the terminal heterocysts seems independent of the filament length because every filament ranging from 3 to 15 cells in length, resulting from sonication of the PAT-1 strain, has at least one heterocyst on one end; most have heterocysts at both ends. A few intercalary heterocysts form in a later stage of growth when the filaments contain >100 cells. In contrast, under the same growth conditions, wild-type filaments differentiate both terminal and internal heterocysts at intervals of ≈10 cells. The heterocysts of strain PAT-1 are morphologically indistinguishable from those of wild type, based on electron microscopy of thin sections (K. Black, personal communication).

Complementation of the Mutation. The strain PAT-1 was complemented by conjugation with a cosmid library of wild-type DNA fragments, selecting for rapid growth on medium lacking combined nitrogen (10). Fifteen cosmids isolated from complemented colonies contained a common 10-kb fragment. One of these cosmids was partially digested with *Sau3*AI, and the resultant fragments of 1–2, 2–4, and 4–14 kb were subcloned into the shuttle vector pCCB110. Complementation of strain PAT-1 by using these three fragment-libraries resulted in the isolation of plasmids from the complemented cells that shared a 1.7-kb region. The complete DNA sequence of plasmids pJHLb-5 and pJHLb-10, carrying the same 2.7-kb insert in opposite orientations, was determined. Fig. 2 shows that this fragment contains a complete open reading frame (ORF) of 1.1 kb and a partial ORF on the 5′ end (data not shown). To confirm that the 1.1-kb ORF was the gene responsible for complementing the mutation, serial deletions from both ends were constructed and used in complementation experiments. The plasmids in which the deletions entered the ORF did not complement the mutation, whereas the plasmids containing the intact ORF did. Thus,

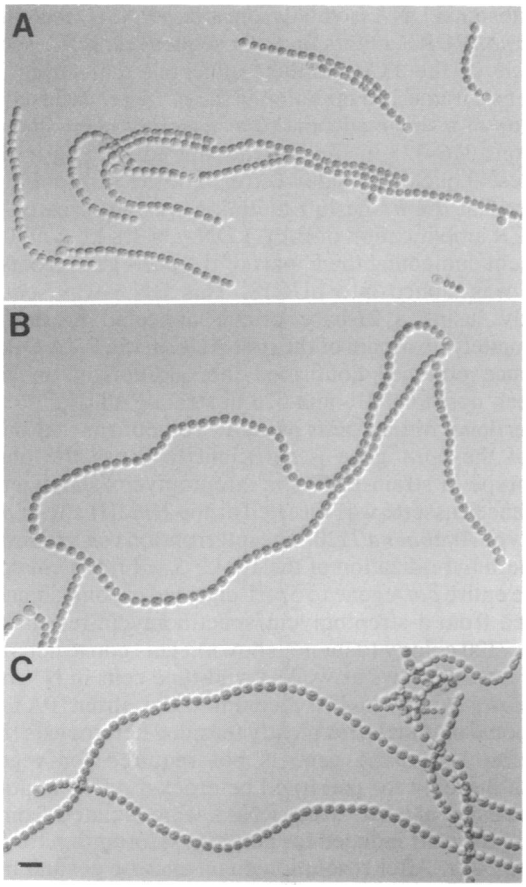

FIG. 1. Differential-interference-contrast micrographs of *Anabaena* 7120 filaments 2 days after transfer to nitrogen-free medium. (*A*) Wild type, in which heterocysts are spaced about 10 cells apart. (*B*) Mutant PAT-1 (note absence of intercalary heterocysts). (*C*) Strain in which the *patA* gene is interrupted by the Ω cassette. This strain appears the same as PAT-1. (Bar = 10 μm.)

the 1.1-kb ORF represents the gene required for proper heterocyst-pattern formation. This gene was named *patA* (for pattern).

Sequence of *patA*. Nested deletions were constructed starting with plasmids pJHLb-5 and pJHLb-10. The complete sequence of both strands of the insert DNA was determined. Fig. 3 presents the DNA sequence of *patA* and part of its upstream sequence. The *patA* gene encodes a protein of 379 amino acids that include 8.8% acidic residues and 12.3%

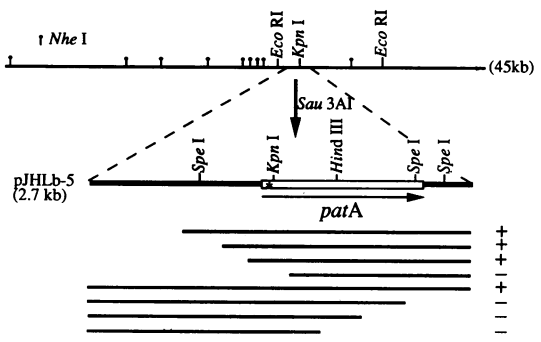

FIG. 2. Physical map of the cosmid containing the *patA* gene, the 2.7-kb fragment that was sequenced, and the deletions used for complementation of the mutation in PAT-1. Horizontal arrow indicates the direction of transcription of the *patA* gene. +/− signs on right indicate whether a plasmid corresponding to the thin horizontal line did or did not complement the mutation in PAT-1. The frameshift in PAT-1 is located at the star (*) within the *patA* box.

```
TCTAGTAATCGGTTGAACAATTTTTCTGTTTGACATCATGAAACTATTCCACTCAGGAGA  60
GTAGCACTGACGGTGTAGAACTCTCCTGAGTGGAAAGATTGGCGTGAAAATAAAAAATAA 120
AGACTAACTGTGTTGACTTAACTGTAAATTGCTGAACTTAGGTAAATATATTCTCAGAAA 180
TACATTTACCTACAGGAAAAGCCTTTTCCCTAATTTTTATTTTAAATCTTGATTTAATAA 240
CAGAGCGATCGCCATGAAAACACTTCCGATTACTAGATACAGATTTTTTCAGAAAATACA 300
            M  K  T  L  P  I  T  R  Y  R  F  F  Q  K  I  Q
                *
ACCCCTATCGTTATTAAAAAAAATAACTGGGAAGACAATCACTGGTTGTCTACAAGTATT 360
 P  L  S  L  L  K  K  I  T  G  K  T  I  T  G  C  L  Q  V  F
TAGTACTTCAGGTACCTGGTCTATATACGTAGAAGAGGGTAAGCTAATTTATGCCTGTTA 420
 S  T  S  G  T  W  S  I  Y  V  E  E  G  K  L  I  Y  A  C  Y
TTCAGAGAGGATGTTTGAACCTCTTTATAGACATTTGGGAAATTTAAGCCCACAAATTGC 480
 S  E  R  M  F  E  P  L  Y  R  H  L  G  N  L  S  P  Q  I  A
AACCCTACCTAAAGAAATTAACGAGCAGTTAAGAGCAATCTTTGAAACAGGAATTGAAAA 540
 T  L  P  K  E  I  N  E  Q  L  R  A  I  F  E  T  G  I  E  N
TCAAGCAATACCCAATCCTGATTATTTAGCTATTTGTTGGTTAGTCAATCAGAAATATAT 600
 Q  A  I  P  N  P  D  Y  L  A  I  C  W  L  V  N  Q  K  Y  I
TAGTTCTTCACAAGCAGCTGTTCTTATAGAACAATTAGCATTAGAAGTAGTAGAATCATT 660
 S  S  S  Q  A  A  V  L  I  E  Q  L  A  L  E  V  V  E  S  F
TCTCATGCTGGAAGAAGGAAGTTATGAATTTATTCCTGAAAGTTTTTTAGATGATTTGCC 720
 L  M  L  E  E  G  S  Y  E  F  I  P  E  S  F  L  D  D  L  P
TAAATTTTGCTATCTGAATGTTCGGTTATTAGTAGAACAGTGTCAACAGCATGGGCGTGT 780
 K  F  C  Y  L  N  V  R  L  L  V  E  Q  C  Q  Q  H  G  R  V
TCCGGAAGCATTCCGCAGAGAAGCTTCCAGTCAAGAGATATCTTCATCTACAGAACATAA 840
 P  E  A  F  R  R  E  A  S  S  Q  E  I  S  S  S  T  E  H  N
CCAGATACCAGTTAACAACAGACGTAGAACCAAATTTACATCTCCTCCTCATACTCAGCC 900
 Q  I  P  V  N  N  R  R  S  T  K  F  T  S  P  P  H  T  Q  P
TAAGCCTGAGCCACGCTTACCGCAAATAAATACTAATAAATCTACAGAATATTCAAAGCG 960
 K  P  E  P  R  L  P  Q  I  N  T  N  K  S  T  E  Y  S  K  R
TTACGCATCTCAACCTAATACTGTCAATCATGGATACTCACAGACATCTGCCACATCTAC 1020
 Y  A  S  Q  P  N  T  V  N  H  G  Y  S  Q  T  S  A  T  S  T
TGATAAAAAAATCTATACAATCTTTTGTATTGATGAGAATCCAATTGTGCTAAATAATAT 1080
 D  K  K  I  Y  T  I  F  C  I  D  E  N  P  I  V  L  N  N  I
TAAAAACTTTTAGATGATCAAATATTTGCAGTAATTGGTGTTACAGATTCTCTAAAAGC 1140
 K  N  F  L  D  D  Q  I  F  A  V  I  G  V  T  D  S  L  K  A
ATTAATGGAAATTCTCTGTACAAAGCCAGATATCATTTTGATCAATGTTGATATGCCTGA 1200
 L  M  E  I  L  C  T  K  P  D  I  I  L  I  N  V  D  M  P  D
TTTAGACGGCTATGAGTTATGTTCTTTATTACGCAAACATTCATATTTTAAAAACACACC 1260
 L  D  G  Y  E  L  C  S  L  L  R  K  H  S  Y  F  K  N  T  P
TGTGATTATGGTGACAGAAAAAGCTGGACTAGTTGATAGAGCGAGAGCAAAGATAGTCAG 1320
 V  I  M  V  T  E  K  A  G  L  V  D  R  A  R  A  K  I  V  R
AGCATCAGGTCACTTAACTAAACCTTTTAATCAAGGTGATTTGCTAAAAGTAATTTTTAA 1380
 A  S  G  H  L  T  K  P  F  N  Q  G  D  L  L  K  V  I  F  K
ACACATTACGTAATATCAACTATTACCAATTACCCATTACCTTCTTAGATGTAGATACAT 1440
 H  I  T  *
```

FIG. 3. Nucleotide sequence of part of the 2.7-kb fragment shown in Fig. 2. The double underlines and wavy underlines in the first five rows indicate inverted repeats. The overlined sequence is complementary to the 3′ end of 16s rRNA. The star (★) indicates the position at which an adenine is inserted to create a frameshift in the PAT-1 strain. The underlined sequences within the coding region are direct repeats that result in repetition of the amino acid sequence FLDD. The residues in boldface type correspond to the conserved D and K residues of the CheY family of regulator proteins.

basic residues. Analysis of the derived protein sequence by the PLOT.A/HYD program (provided by P. Markiewicz, University of California, Los Angeles) using the Kyte/Doolittle parameter set (22) revealed that the overall hydrophobicity of the protein can be divided into three distinguishable regions. Although the region between residues 168 and 259 is hydrophilic, the amino-terminal and the carboxyl-terminal parts alternate between hydrophobic and hydrophilic segments (Fig. 4). Analysis of the same sequence by PLOT.A/KAS (a flexibility predictor, also provided by P. Markiewicz) indicated that the middle part of the protein (residues 168–259) is more flexible than its flanking sequences (data not shown).

The nucleotide sequence contains three pairs of repeated sequences in the coding region and the upstream region. One pair is a direct repeat of 12 nucleotides (TTTTTAGATGAT) in the coding region. A second pair is an inverted repeat of 16 nucleotides (TTCCACTCAGGAGAGT) in the upstream region, and the third pair is also an inverted repeat of 12 nucleotides (TCTAGTAATCGG) with one-half in the upstream region and the other half in the amino terminus of the *patA*-coding region. No function has been ascribed to these repeats so far. A putative ribosome-binding site is located three bases upstream of the first ATG codon.

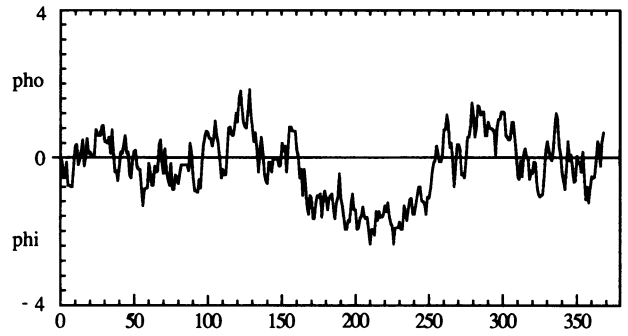

FIG. 4. Hydropathy plot of the PatA protein. The ordinate shows hydrophobicity calculated according to Kyte and Doolittle (22) with a window of seven residues. PatA sequence similarity to the CheY protein of *E. coli* begins at residue 256 (see Fig. 5), just beyond the junction between the central hydrophilic region and the carboxyl-terminal third of PatA. pho, Hydrophobicity; phi, hydrophilicity.

A Region of PatA Is Homologous to CheY. The predicted amino acid sequence of PatA was compared with the patterns derived from the sequences of homologous protein families in the European Molecular Biology Laboratory Swiss-Prot protein sequence data base (release 13) by using the PLSEARCH program (release no. 4, developed by R. F. Smith and T. F. Smith, Molecular Biology Computer Research Resource, 1990) (23). This search identified significant similarity between the carboxyl-terminal region of PatA and the CheY protein of *E. coli* (13). The carboxyl-terminal fragment of PatA was subsequently compared with the protein sequences in GenBank (release no. 70) with the computer program FASTA (24). High similarity was found between the carboxyl-terminal region of PatA and the amino-terminal regions of a group of proteins involved in bacterial sensory-transduction systems. No significant similarity was found between the amino-terminal region of PatA and any other proteins in the data base. Fig. 5 shows pairwise alignments of PatA with CheY and OmpR with the computer program CLUSTAL. CheY, for which the x-ray crystal structure has been solved, is a small regulator protein of 129 amino acids; it is composed of a central core of five parallel β-strands and an outer layer of five α-helices (26). Residues Asp-12, Asp-13, Asp-57, and Lys-109 of CheY, which all appear at the carboxyl end of the β-sheets, are conserved at the corresponding positions in most of the response regulator proteins (27) and in PatA.

Strain PAT-1 Contains a Frameshift Mutation in *patA*. Because no Tn*5* was detected in the chromosomal DNA of strain PAT-1, the nature of the mutation was of interest. We made two primers and amplified a 1.8-kb fragment of the chromosomal DNA isolated from strain PAT-1, including the entire *patA* ORF and its flanking sequences. DNA sequence analysis of the PCR product indicated that strain PAT-1 contains a frameshift mutation in the *patA* gene caused by the addition of a single adenine to a stretch of eight adenines at positions 316–323 in Fig. 3. The altered reading frame terminates 40 nucleotides downstream of the shift point. To be certain that the frameshift mutation was not created during the PCR amplification of PAT-1 DNA, a 3.8-kb *Nhe* I–*Kpn* I fragment containing the 5′ part of the *patA* gene from PAT-1 DNA was cloned into pUC18. This DNA was sequenced directly, using a 21-base primer annealed to the region immediately upstream of the first AUG in the PatA ORF. The sequence obtained confirmed the addition of an adenine between positions 316 and 323 in strain PAT-1.

```
Ec CheY  MADKELK-FLVVDDFSTMRRIVRNLLKELGFNNVEEAEDGVDAL
          *||**  ** *|* *** * **|*|**  |* | ***|***||
An PatA  >TDKKIYTIFCIDENPIVLNNIKNFLDDQIFA-VIGVTDSLKAL
           **  |*|* *|**  * * ****|**| |* | *|***  *
St OmpR  MQEN--YKILVVDDDMRLRALLERYLTEQGFQ-VRSVANA--EQ

Ec CheY  NKL--QAGGFGFIISDWNMPNMDGLELLKTIRADSAMSALPVLM
          **     ** *|  |   |**|| || * *| *| *** ||*|
An PatA  MEILCTKPDIILINVDM--PDLDGYELCSLLRKHSYFKNTPVIM
         |* | |* ** |* *|*  |* || **|* ||**|  *  |*||
St OmpR  MDRLLTRESFHLMVLDLMLPGEDGLSICRRLRSQS--NPMPIIM

Ec CheY  VTAEAKKENIIAAAQAGASGYVVKPFTAATLEEKLNKIFEKLGM
         ||**|   *      * |||***|||****|   |* ||****
An PatA  VTEKAGLVDRARRKIVRASGHLTKPFNQGDL---LKVIFKHIT
         ||*|** |||     * |*****||||* *|   ** **** *
St OmpR  VTAKGEEVDRIVGLEIGADDYIPKPFNPRELLARIRPVLRRQA>
```

FIG. 5. Comparison of the amino acid sequences of *Anabaena* PatA (An PatA) (residues 256–379), all of *E. coli* CheY (Ec CheY) (13), and the amino-terminal region of OmpR from *Salmonella typhimurium* (St OmpR) (25). Identical residues are indicated by vertical bars, similar residues are shown by stars, and the critical aspartate and lysine residues are in boldface type. The analysis was done by using the computer program CLUSTAL with the standard Pam 250 matrix for proteins and gap penalty of 10.

Insertional Mutagenesis of *patA*. To confirm that inactivation of the *patA* gene is sufficient to cause the observed phenotype in strain PAT-1, a streptomycin/spectinomycin-resistance cassette was inserted at the *Hin*dIII site in *patA* in wild-type *Anabaena* 7120. The interruption was confirmed by Southern hybridization of the 1.4-kb *Spe* I fragment containing the entire *patA* gene to *Spe* I-digested chromosomal DNA isolated from a streptomycin/spectinomycin-resistant *Anabaena* 7120 colony (data not shown). This constructed insertional mutant grows as well as wild-type cells in N^+ medium but grows poorly in N^- medium. Like strain PAT-1, the insertional mutant forms mostly terminal heterocysts (Fig. 1). Therefore, the *patA* gene is not required for vegetative growth but only for patterned heterocyst differentiation.

Expression of *patA*. Total RNAs were isolated from *Anabaena* 7120 cells induced for heterocyst formation for 0, 3, 6, 12, and 24 hr. After resolution on an agarose gel and transfer to a nylon membrane, the RNAs were probed with the 1.4-kb *Spe* I fragment containing *patA*. Fig. 6 shows that a 1.4-kb band hybridizing to the probe was seen in all the RNA samples. Between 3 and 6 hr the abundance of the 1.4-kb band increases slightly. This pattern of expression resembles that of the *hetR* gene (11). The blot in Fig. 6 had to be exposed to film for >2 weeks, suggesting that the *patA* gene transcript is much less abundant than the *hetR* gene transcript, which appears in exposures of <1 day (11). The size of the mRNA indicates that *patA* is not cotranscribed with any other gene.

Suppression of the Extracopy *hetR* Phenotype. When the wild-type *hetR* gene is introduced into wild-type *Anabaena* 7120 on a plasmid, the pattern of heterocyst differentiation is severely perturbed (11). Filaments differentiate heterocysts in medium containing NH_4^+ or NO_3^-. Cultures induced by

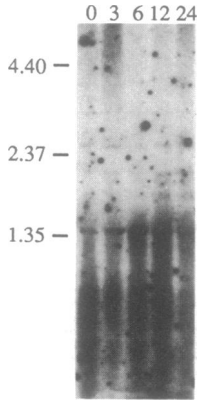

FIG. 6. Expression of the *patA* gene. The RNA blot shown contained total RNA (40 μg in each lane) prepared from wild-type *Anabaena* 7120 starved for combined nitrogen for 0, 3, 6, 12, or 24 hr. The probe was a 1.4-kb *Spe* I fragment containing *patA* (see Fig. 2). The 1.4-kb *patA* mRNA is seen in all lanes, increasing in abundance after 3 hr of induction. The film was exposed for 16 days.

nitrogen starvation differentiate strings of heterocysts containing two, three, or more cells (11). When the same *hetR*-containing plasmid was conjugated into the *patA* mutant strain, the resulting exconjugants retained the PAT-1 phenotype, with single heterocysts at the ends of filaments, only under inducing conditions.

DISCUSSION

The *patA* Mutation in Strain PAT-1. The original mutation in strain PAT-1 is a frameshift resulting from addition of an adenine to a string of adenines. This result appears to be a spontaneous mutation of the type arising from slippage during replication rather than an insertion due to movement of transposon Tn*5* in and out of the chromosome. Thus, the mutation is probably not related to the initial selection for neomycin resistance. However, the Fix^- phenotype is due to the *patA* mutation because it is complemented only by the complete *patA* gene, and the phenotype can be duplicated by interruption of the *patA* gene in a wild-type strain. A mutant with a similar phenotype was described in a different *Anabaena* strain many years ago, but molecular methods for genetic analysis were not available then, and the mutant has not been studied further (9).

Function of the *patA* Gene. The *patA* gene product is required for the differentiation of intercalary heterocysts but not for terminal heterocysts. The terminal heterocysts that develop in the PAT-1 strain are morphologically normal at both the light microscope and electron microscope level, and they are capable of nitrogen fixation. A *patA* mutant grows slowly on N^- medium because the heterocysts cannot supply the need of the entire filament for fixed nitrogen. Although the *patA* gene is transcribed during growth on fixed nitrogen, it is not needed for that growth. After a switch to nitrogen starvation conditions, it is transcribed more abundantly. The *patA* gene transcript is at least an order of magnitude less abundant than the *hetR* gene transcript at all times tested. The presence of both transcripts in vegetative cells is consistent with the participation of their products in an environment-sensing regulatory circuit. In enterobacteriaceae, for example, both *ntrB* and *ntrC* are transcribed at a basal level under nitrogen-replete conditions, but the abundance of both gene transcripts increases during nitrogen step-down as a consequence of readthrough from the *glnA* promoter that they activate (28).

Given the sequence similarity between PatA and the CheY family of regulatory proteins, we might suggest that PatA is a transcription activator controlled by phosphorylation of an aspartate residue (29). HetR might also be a component of the environment-sensing system. But these two proteins cannot be obligatorily linked to the transcription of genes required for heterocyst differentiation because the *patA* mutation permits terminal heterocyst differentiation—i.e., the *hetR* gene functions in those terminal cells. On the other hand, the *patA* mutation prevents the consequences of the overexpression of *hetR* in intercalary cells in the extracopy *hetR* strain. These observations suggest that PatA is required to interpret the environment only in intercalary cells.

The PatA Protein. The carboxyl-terminal third of the PatA protein is similar in sequence to CheY, including the three conserved aspartate residues that constitute the acid pocket, which forms the phosphorylation site of CheY, and one lysine residue (13, 26). Most of the response regulator proteins of the two-component systems have their phosphorylation sites in the amino-terminal part of the protein (27). These proteins, except for CheB (a substrate for the CheA kinase that also phosphorylates CheY), activate transcription when phosphorylated. CheY is much smaller than the transcription activators, consisting of only the conserved phosphorylation domain. CheY activates the flagellar motor, rather than transcription (14). PatA also differs from the transcription activators, having its putative phosphorylation site in its carboxyl-terminal third.

We thank W. Buikema, C. Halling, P. Gornicki, C. Bauer, and K. Black for helpful discussions and for critical reading of the manuscript. This work was supported by Research Grants GM40685 and GM21823 from the National Institutes of Health.

1. Haselkorn, R. & Buikema, W. J. (1992) in *Biological Nitrogen Fixation*, eds. Stacey, G., Burris, R. H. & Evans, H. J. (Chapman & Hall, New York), pp. 166–190.
2. Walsby, A. E. (1985) *Proc. R. Soc. London B* **226**, 345–366.
3. Wilcox, M., Mitchison, G. J. & Smith, R. J. (1973) *J. Cell Sci.* **12**, 707–723.
4. Wolk, C. P. & Quine, M. P. (1975) *Dev. Biol.* **46**, 370–382.
5. Mitchison, G. J. & Wilcox, M. (1973) *Nature (London) New Biol.* **246**, 229–233.
6. Haselkorn, R. (1978) *Annu. Rev. Plant Physiol.* **29**, 319–344.
7. Wolk, C. P. (1967) *Proc. Natl. Acad. Sci. USA* **57**, 1246–1251.
8. Wolk, C. P. (1989) *Plant Syst. Evol.* **164**, 27–31.
9. Wilcox, M., Mitchison, G. J. & Smith, R. J. (1975) *Arch. Microbiol.* **103**, 219–223.
10. Buikema, W. J. & Haselkorn, R. (1991) *J. Bacteriol.* **173**, 1879–1885.
11. Buikema, W. J. & Haselkorn, R. (1991) *Genes Dev.* **5**, 321–330.
12. Wolk, C. P. (1991) *Curr. Opin. Genet. Dev.* **1**, 336–341.
13. Matsumura, P., Rydel, J. J., Linzmeier, R. & Vacante, D. (1984) *J. Bacteriol.* **160**, 36–41.
14. Bourret, R. B., Borkovich, K. A. & Simon, M. I. (1991) *Annu. Rev. Biochem.* **60**, 401–441.
15. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
16. Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. & Stanier, R. Y. (1979) *J. Gen. Microbiol.* **111**, 1–61.
17. Borthakur, D. & Haselkorn, R. (1989) *J. Bacteriol.* **171**, 5759–5761.
18. Golden, J. W. & Wiest, D. R. (1988) *Science* **242**, 1421–1423.
19. Cai, Y. & Wolk, C. P. (1990) *J. Bacteriol.* **172**, 3138–3145.
20. Wolk, C. P., Vonshak, A., Kehoe, P. & Elhai, J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1561–1565.
21. Elhai, J. & Wolk, C. P. (1988) *Methods Enzymol.* **167**, 747–754.
22. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132.
23. Smith, R. F. & Smith, T. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 118–122.
24. Pearson, W. R. & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 2444–2448.
25. Liljeström, P., Laamenen, I. & Palva, E. T. (1988) *J. Mol. Biol.* **201**, 663–673.
26. Stock, A. M., Mottonen, J. M., Stock, J. B. & Schutt, C. E. (1989) *Nature (London)* **337**, 745–749.
27. Stock, J. B., Stock, A. M. & Mottonen, J. M. (1990) *Nature (London)* **344**, 395–400.
28. Ninfa, A. J. & Magasanik, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5909–5913.
29. Stock, J. B., Ninfa, A. J. & Stock, A. M. (1989) *Microbiol. Rev.* **53**, 450–490.